From the Kochen-Specker theorem to noncontextuality inequalities without assuming determinism

Ravi Kunjwal
Optics & Quantum Information Group, The Institute of Mathematical Sciences, C.I.T Campus, Taramani, Chennai 600 113, India

Robert W. Spekkens
Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario Canada N2L 2Y5

The Kochen-Specker theorem demonstrates that it is not possible to reproduce the predictions of quantum theory in terms of a hidden variable model where the hidden variables assign a value to every projector deterministically and noncontextually. A noncontextual value-assignment to a projector is one that does not depend on which other projectors—the context—are measured together with it. Using a generalization of the notion of noncontextuality that applies to both measurements and preparations, we propose a scheme for deriving inequalities that test whether a given set of experimental statistics is consistent with a noncontextual model. Unlike previous inequalities inspired by the Kochen-Specker theorem, we do not assume that the value-assignments are deterministic and therefore in the face of a violation of our inequality, the possibility of salvaging noncontextuality by abandoning determinism is no longer an option. Our approach is operational in the sense that it does not presume quantum theory: a violation of our inequality implies the impossibility of a noncontextual model for *any* operational theory that can account for the experimental observations, including any successor to quantum theory.

PACS numbers: 03.65.Ta, 03.65.Ud

Although measurements in quantum theory cannot, in general, be implemented simultaneously, one can still ask whether the outcomes of such incompatible measurements might be simultaneously well-defined within some deeper theory. To formalize this deeper theory we use the framework of *ontological models* [1] which generalizes the notion of a hidden variable model. Contrary to naïve impressions, it is possible to find models of this sort that reproduce quantum predictions. Problems only arise if one makes additional assumptions about the model. The Kochen-Specker theorem [2] famously derives a contradiction from an assumption we term *KS-noncontextuality*. Consider a set of quantum measurements, each represented by an orthonormal basis, such that some rays are common to more than one basis. It is assumed that every *ontic state*—a complete specification of the properties of the system, including values of hidden variables—assigns a definite value to each ray, 0 or 1, *regardless of the basis (i.e. context) in which the ray appears*. If a ray is assigned the value 1 (0) by an ontic state λ, the measurement outcome associated with that ray is predicted to occur with probability 1 (0) when any measurement including the ray is implemented on the system in ontic state λ. It follows that for every basis, precisely one ray must be assigned the value 1 and the others the value 0.

The assumption that the ontic state assigns a deterministic outcome to each measurement is the greatest shortcoming of the Kochen-Specker theorem. Recall that determinism is not an assumption of Bell's theorem [3, 4]. This is evident from derivations of the Clauser-Horne-Shimony-Holt inequality [5]. Even in Bell's original 1964 article [3], where deterministic assignments play an important role, determinism is not assumed but rather *derived* from local causality and the fact that quantum theory predicts perfect correlations if the same observable is measured on the two parts of a maximally entangled state (an argument from Einstein, Podolsky and Rosen [6] that Bell simply recycled [7]). It was shown in Ref. [8] that one can make a similar argument about determinism in noncontextual models: rather than assuming it, one can derive it from a generalized notion of noncontextuality and from two facts about quantum theory: (i) the outcome of a measurement of some observable is perfectly predictable whenever the preceding preparation is of an eigenstate of that observable, and (ii) the indistinguishability, relative to all quantum measurements, of different convex decompositions of the completely mixed state into pure states.

Hence, in any proof of the Kochen-Specker theorem one can replace the assumption of determinism with the generalized notion of noncontextuality and the quantum prediction of perfect predictability. If perfect predictability is indeed observed, then in the face of the resulting contradiction, one must give up on noncontextuality. This contrasts with earlier proofs where one could always salvage the generalized notion of noncontextuality by abandoning determinism.

Of course, no real experiment ever yields *perfect* predictability, so this manner of ruling out noncontextuality is not robust to experimental error. Following ideas introduced in recent work [9], we show how to contend with the lack of perfect predictability of measurements and derive an experimentally-robust noncontextuality inequality for any uncolourability proof of the Kochen-Specker theorem.

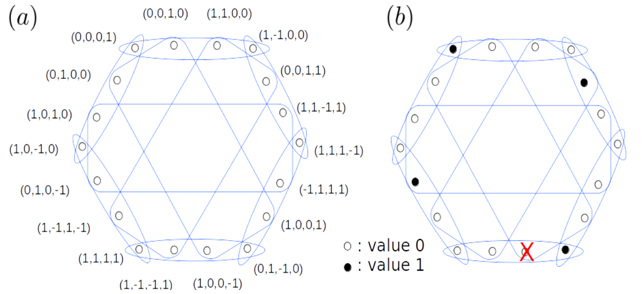

FIG. 1. Each of the 18 rays is depicted by a node, and the 9 orthonormal bases are depicted by 9 edges, each a loop encircling 4 nodes. There is no noncontextual assignment of 0s and 1s to these nodes such that for every edge precisely one node receives the value 1. For instance, we have depicted a noncontextual assignment of 0s and 1s to 17 of the rays, which cannot be completed to an assignment to all 18 rays because neither value (0 or 1) can be assigned to the remaining ray (marked by X): while one basis in which it appears requires it to take the value 0, the other requires the value 1.

Review of the Kochen-Specker theorem. The original proof of the KS theorem required 117 rays in a 3d Hilbert space [2]. We use the much simpler proof in Ref. [10] as our illustrative example. It involves a 4d Hilbert space and 18 rays that appear in 9 orthonormal bases, each ray appearing in two bases. One can visualize this as a hypergraph with nodes representing the rays and edges representing orthonormal bases (Fig. 1(a)). There is no 0-1 assignment to these rays that respects KS-noncontextuality: the hypergraph is *uncolourable* (Fig. 1(b)). Of course, if the value assigned to a ray were allowed to be 0 in one basis and 1 in the other (a KS-contextual value assignment) then one could evade the contradiction.

Is it possible to test the possibility of a KS-noncontextual ontological model experimentally? One view is that the Kochen-Specker theorem is not amenable to an experimental test. It merely constrains the possibilities for *interpreting* the quantum formalism [11, 12]. However, this answer is clearly inadequate. One *can* and *should* ask: what is the minimal set of operational predictions of quantum theory that need to be experimentally verified in order to show that it does not admit of a noncontextual model?

We show that this minimal set is a far cry from the whole of quantum theory and is therefore consistent with many other possible operational theories. As such, the no-go result we derive shows that none of these theories admit of a noncontextual model. Furthermore, if this set of predictions is corroborated by experiment, then this implies that any future theory of physics that might replace quantum theory also fails to admit of a noncontextual model.

We begin with some definitions. An *operational theory* is a triple $(\mathcal{P}, \mathcal{M}, p)$ where $\mathcal{P}$ is a set of preparations, $\mathcal{M}$ is a set of measurements, and p specifies, for every pair of preparation and measurement, the probability distribution over outcomes for that measurement if it is implemented on that preparation. Specifically, if we denote the set of outcomes of measurement M by $\mathcal{K}_M$, then $\forall P \in \mathcal{P}$, $\forall M \in \mathcal{M}$, p is a function of the form $p(\cdot|P, M) : \mathcal{K}_M \to [0, 1]$.

An *ontological model* of an operational theory $(\mathcal{P}, \mathcal{M}, p)$ is a triple (Λ, μ, ξ), where Λ denotes a space of possible ontic states for the physical system (here presumed to be discrete), where μ specifies a probability distribution over the ontic states for every preparation procedure, that is, $\forall P \in \mathcal{P}$, $\mu(\cdot|P) : \Lambda \to [0, 1]$, such that $\sum_{\lambda\in\Lambda} \mu(\lambda|P) = 1$, and where ξ specifies, for every measurement, the conditional probability of obtaining a given outcome if the system is in a particular ontic state, that is, $\forall M \in \mathcal{M}$, $\xi(k|M, \cdot) : \Lambda \to [0, 1]$, such that $\sum_{k\in\mathcal{K}_M} \xi(k|M, \lambda) = 1$. In order for the ontological model to reproduce the statistical predictions of the operational theory, it must be the case that

$$p(k|P, M) = \sum_{\lambda\in\Lambda} \xi(k|M, \lambda)\mu(\lambda|P) \qquad (1)$$

for all $P \in \mathcal{P}$, and $M \in \mathcal{M}$.

We denote the event of obtaining outcome k of measurement M by $[k|M]$. If $[k|M]$ is assigned a deterministic outcome by every ontic state in the ontological model, i.e., if $\xi(k|M, \cdot) : \Lambda \to \{0, 1\}$, then it is said to be *outcome-deterministic* in that model, and if this holds for all k, then M is also said to be outcome-deterministic.

We explain how to derive an experimental test of noncontextuality using a sequence of four refinements on the standard account of the KS theorem:

Operationalizing the notion of KS-noncontextuality. In a KS-noncontextual model of operational quantum theory, the value (0 or 1) assigned to the event $[k|M]$ by λ is the same as the value assigned to the event $[k'|M']$ whenever these two events are represented by the same ray of Hilbert space (here, we are assuming that M and M' are maximal projective measurements). We get to the crux of the notion of KS-noncontextuality, therefore, by describing the *operational grounds* for associating the same ray to $[k|M]$ as is associated to $[k'|M']$. Letting $\Pi_{k|M}$ and $\Pi_{k'|M'}$ represent the corresponding rank-1 projectors, the grounds for concluding that $\Pi_{k|M} = \Pi_{k'|M'}$ are that $\mathrm{tr}(\rho\Pi_{k|M}) = \mathrm{tr}(\rho\Pi_{k'|M'})$ for an appropriate set of density operators ρ. It is clearly sufficient for the equality to hold for the set of *all* density operators, but it is also sufficient to have equality for certain smaller sets of density operators, namely, those *complete for measurement tomography*, or simply *tomographically complete*.

What then should the operational grounds be for assigning the same value to $[k|M]$ and $[k'|M']$ in a general operational theory, where preparations are not represented by density operators? The answer, clearly, is that the event $[k|M]$ *occurs with the same probability* as the event $[k'|M']$ for *all* preparation procedures of the

system,

$$p(k|M,P) = p(k'|M',P) \ \text{ for all } \ P \in \mathcal{P}, \qquad (2)$$

or equivalently, if this holds for a subset of $\mathcal{P}$ that is tomographically complete. In this case, we shall say that $[k|M]$ and $[k'|M']$ are *operationally equivalent*, and denote this as $[k|M] \simeq [k'|M']$. We can therefore define a notion of KS-noncontextuality for any operational theory as follows: an ontological model (Λ, μ, ξ) of an operational theory $(\mathcal{P}, \mathcal{M}, p)$ is KS-noncontextual if (i) operational equivalence of events implies equivalent representations in the model, i.e., $[k|M] \simeq [k'|M'] \Rightarrow \xi(k|M,\lambda) = \xi(k'|M',\lambda)$ for all $\lambda \in \Lambda$, and (ii) the model is outcome-deterministic, $\xi(k|M,\cdot) : \Lambda \to \{0,1\}$.

The operational equivalences among the measurements that are relevant for the 18 ray proof of the KS theorem depicted in Fig. 1(a) are made explicit in Fig. 2(a), where every measurement event $[k|M]$ is represented by a distinct node, and a novel type of edge between nodes specifies when two events are operationally equivalent. This representation affords a nice way of depicting contextual value assignments, such as in Fig. 2(b). It follows that *any* operational theory that admits of nine four-outcome measurements that satisfy the operational equivalence relations depicted in Fig. 2(a) fails to admit of a KS-noncontextual model.

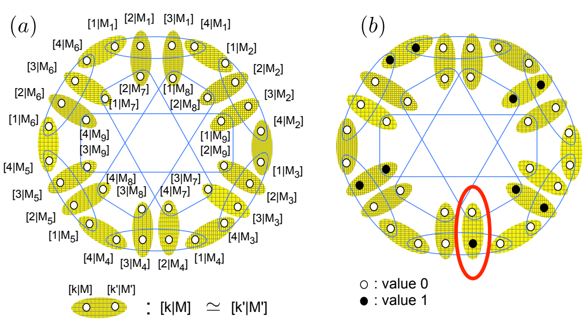

FIG. 2. (a) Nine four-outcome measurements. A blue loop encircling a set of nodes implies that these nodes denote outcomes of a single measurement. A yellow hashed region enclosing a set of nodes implies that the corresponding events are operationally equivalent. (b) A depiction of the fact that there is no outcome-deterministic noncontextual assignment of values in $\{0,1\}$ to the measurements. The depicted value-assignment breaks the assumption of noncontextuality for the pair of highlighted nodes

Defining a notion of noncontextuality without outcome determinism. The essence of noncontextuality is that context-independence at the operational level should imply context-independence at the ontological level. The operationalized version of KS-noncontextuality commits one to more than this, however, because it makes an additional assumption about *what sort of thing* should be independent of context at the ontological level, namely, a deterministic assignment of an outcome. However, one can equally well assume that the ontic state merely assigns a probability distribution over outcomes, and take *this distribution* to be the thing independent of the context. In Ref. [8], this revised notion of noncontextuality was termed *measurement noncontextuality*:

> *Measurement noncontextuality* is satisfied by an ontological model (Λ, μ, ξ) of an operational theory $(\mathcal{P}, \mathcal{M}, p)$ if $[k|M] \simeq [k'|M']$ implies $\xi(k|M,\lambda) = \xi(k'|M',\lambda)$ for all $\lambda \in \Lambda$.

Here, $\xi(k|M,\cdot) \in [0,1]$ (and not merely $\{0,1\}$). Outcome determinism is not assumed.

Justifying outcome determinism for perfectly predictable measurements. Outcome determinism can, however, be justified sometimes if one assumes a notion of noncontextuality for *preparations* [8]. First, a definition: P and P' are said to be operationally equivalent, denoted $P \simeq P'$, if for every measurement event $[k|M]$, P assigns the same probability to this event as P' does, that is,

$$p(k|M,P) = p(k|M,P') \ \text{ for all } \ k \in \mathcal{K}_M, \ \text{ for all } \ M \in \mathcal{M}. \qquad (3)$$

A preparation-noncontextual ontological model is then defined as follows:

> *Preparation noncontextuality* is satisfied by an ontological model (Λ, μ, ξ) of an operational theory $(\mathcal{P}, \mathcal{M}, p)$ if $P \simeq P'$ implies $\mu(\lambda|P) = \mu(\lambda|P')$ for all $\lambda \in \Lambda$.

Insofar as both measurement and preparation noncontextuality are instances of operational equivalence implying ontological equivalence, it is most natural to assume *both*, that is, to assume *universal noncontextuality*.

It was shown in Ref. [8] that in a preparation-noncontextual model of quantum theory, all projective measurements must be represented outcome-deterministically. Here, we provide a version of this argument for the 18 ray construction.

Suppose that one has experimentally identified thirty-six preparation procedures organized into nine ensembles of four each, $\{P_{i,k} : i \in \{1,\ldots,9\}, k \in \{1,\ldots,4\}\}$, such that for all i, measurement M_i on preparation $P_{i,k}$ yields the kth outcome with certainty,

$$\forall i, \forall k : p(k|M_i, P_{i,k}) = 1. \qquad (4)$$

We call this property *perfect correlation*. In quantum theory, it suffices to let $P_{i,k}$ be the preparation associated with the pure state corresponding to the kth element of the ith measurement basis.

Define the effective preparation $P_i^{(\mathrm{ave})}$ as the procedure obtained by sampling k uniformly at random and then implementing $P_{i,k}$. We now suppose that one has experimentally verified the operational equivalence relations

$$P_i^{(\mathrm{ave})} \simeq P_{i'}^{(\mathrm{ave})} \ \text{ for all } i, i' \in \{1,\ldots,9\}. \qquad (5)$$

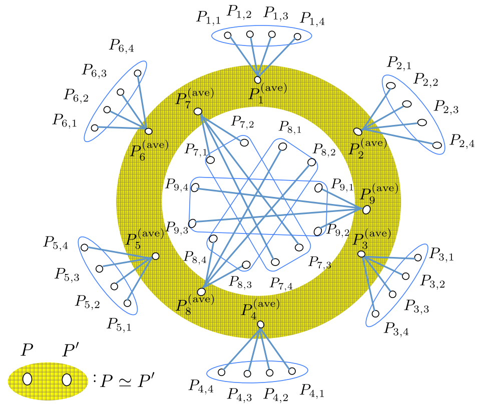

FIG. 3. 36 preparation procedures organized into nine ensembles of four each. A node at the end of a set of lines emanating from the elements of an ensemble represents the effective preparation procedure achieved by sampling uniformly from the ensemble. A yellow region encircling a set of nodes implies that these preparations are operationally equivalent.

These equivalences are depicted in Fig. 3. They hold in our quantum example because the $P_i^{(\mathrm{ave})}$ simply correspond to different ways of preparing the completely mixed state.

Given Eq. (5) and the assumption of preparation noncontextuality, there is a single distribution over Λ, denoted $\nu(\lambda)$, such that

$$\mu(\lambda|P_i^{(\mathrm{ave})}) = \nu(\lambda) \;\; \text{for all } i \in \{1,\ldots,9\}. \tag{6}$$

Given the definition of $P_i^{(\mathrm{ave})}$, it follows that

$$\frac{1}{4}\sum_k \mu(\lambda|P_{i,k}) = \nu(\lambda) \;\; \text{for all } i \in \{1,\ldots,9\}. \tag{7}$$

Furthermore, recalling Eq. (1), for the ontological model to reproduce Eq. (4), we must have

$$\forall i, \forall k : \sum_\lambda \xi(k|M_i,\lambda)\mu(\lambda|P_{i,k}) = 1. \tag{8}$$

Because every λ in the support of $\nu(\lambda)$ appears in the support of $\mu(\lambda|P_{i,k})$ for some k, it follows that if $\xi(k|M_i,\lambda)$ had an indeterministic response on any such λ, we would have a contradiction with Eq. (8). Consequently, for all i and k, the measurement event $[k|M_i]$ must be outcome-deterministic for all λ in the support of $\nu(\lambda)$.

To summarize then, if one has experimentally verified the operational equivalences depicted in Figs. 2(a) and 3 and the measurement statistics described in Eq. (4), then universal noncontextuality implies that the value assignments to measurement events should be deterministic and noncontextual, hence KS-noncontextual, and we obtain a contradiction in the usual manner. The argument can be summarized thus

$$\begin{aligned}&\text{universal noncontextuality} + \text{operational equivalences}\\ &\qquad + \text{perfect correlation} \rightarrow \text{contradiction}.\end{aligned} \tag{9}$$

Contending with the lack of perfect predictability in real experiments. In real experiments, the ideal of perfect correlation described by Eq. (4) is never achieved, so we cannot derive a contradiction from it. However, Eq. (9) is logically equivalent to the following inference:

$$\begin{aligned}&\text{universal noncontextuality} + \text{operational equivalences}\\ &\qquad \rightarrow \text{failure of perfect correlation}.\end{aligned} \tag{10}$$

This means that the amount of correlation, averaged over all i and k, will necessarily be bounded away from 1. It is this bound that is the operational noncontextuality inequality. For the 18 ray example, we prove that

$$A \equiv \frac{1}{36}\sum_{i=1}^{9}\sum_{k=1}^{4} p(k|M_i,P_{i,k}) \leq \frac{5}{6}. \tag{11}$$

To test the assumption of noncontextuality, therefore, one must measure the correlation $p(k|M_i,P_{i,k})$ for all i and k, but one must also verify that the operational equivalences depicted in Figs. 2(a) and 3 hold, because only in this case does the assumption of noncontextuality imply that the inequality (11) should hold.

We now outline how the bound in Eq. (11) is obtained. First, we use Eq. (1) to express A in terms of $\xi(k|M_i,\lambda)$ and $\mu(\lambda|P_{i,k})$. Defining the *max-predictability* of a measurement M given an ontic state λ by

$$\zeta(M,\lambda) \equiv \max_{k' \in \mathcal{K}_M} \xi(k'|M,\lambda), \tag{12}$$

we deduce that

$$\begin{aligned} A &\leq \sum_\lambda \left(\frac{1}{9}\sum_i \zeta(M_i,\lambda) \left[\frac{1}{4}\sum_k \mu(\lambda|P_{i,k}) \right] \right) \\ &= \sum_\lambda \left(\frac{1}{9}\sum_i \zeta(M_i,\lambda) \right) \nu(\lambda) \\ &\leq \max_\lambda \left(\frac{1}{9}\sum_i \zeta(M_i,\lambda) \right), \end{aligned} \tag{13}$$

where we have used Eq. (7).

The measurements can have indeterministic responses, $\xi(k|M,\cdot) : \Lambda \rightarrow [0,1]$, but measurement noncontextuality implies that $\xi(k|M_i,\lambda) = \xi(k'|M_{i'},\lambda)$ for the operationally equivalent pairs $\{[k|M_i],[k'|M_{i'}]\}$. There are many such assignments. Every unit-trace positive operator, for instance, specifies an indeterministic noncontextual assignment via the Born rule, and there are other, nonquantum assignments as well, such as the one depicted in Fig. 4. Consider the average max-predictability achieved by the assignment of Fig 4. Here, six measurements have max-predictability 1, while three have max-predictability $\frac{1}{2}$. This implies that $\frac{1}{9}\sum_i \zeta(M_i,\lambda) = \frac{1}{9}(6\cdot 1 + 3\cdot\frac{1}{2}) = \frac{5}{6}$. As we demonstrate in Appendix A, no ontic state has a higher average max-predictability than that of Fig. 4, so that $\max_\lambda \left(\frac{1}{9}\sum_i \zeta(M_i,\lambda)\right) \leq \frac{5}{6}$, thereby establishing the noncontextual bound on A. The

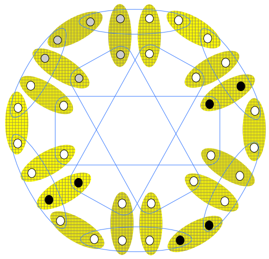

FIG. 4. Example of a noncontextual outcome-indeterministic assignment to the measurements.

logical limit for the value of A is 1, so the noncontextual bound of $\frac{5}{6}$ is nontrivial. The quantum realization of the 18 ray construction achieves $A = 1$.

Note that if an experiment fails to suppress noise sufficiently, then it may not succeed in violating our noncontextuality inequality. This simple criterion of operational meaningfulness fails for previous attempts at deriving noncontextuality inequalities [13], a point we discuss further in Appendices B and C. Although we have used the 18 ray uncolourable set of Ref. [10] as an example, the scheme described can be used to turn any proof of the Kochen-Specker theorem based on an uncolourable set into an experimental inequality. An issue we haven't addressed is that in practice no two measurement events are assigned *exactly* the same probability by each of a tomographically complete set of preparations, nor do any two preparations assign *exactly* the same probability distribution over outcomes to each of a tomographically complete set of measurements. The solution to this problem is described in related work [9, 14]. A question that remains is: how does one accumulate evidence that a given set of measurements or preparations is indeed tomographically complete? This question represents the new frontier in the project of devising strict experimental tests of the assumption of noncontextuality.

Acknowledgments: RK thanks the Perimeter Institute and the Institute of Mathematical Sciences for supporting his visit during the course of this work. This project was made possible in part through the support of a grant from the John Templeton Foundation. Research at Perimeter Institute is supported by the Government of Canada through Industry Canada and by the Province of Ontario through the Ministry of Economic Development and Innovation.

[1] The ontological models framework has the advantage of not prejudicing the question of whether any of the variables remain unkown (i.e. hidden) to one who knows the preparation procedure. For an overview, see N. Harrigan and R. W. Spekkens, Einstein, Incompleteness, and the Epistemic View of Quantum States, Found. Phys. **40**, 125 (2010).

[2] S. Kochen and E. P. Specker, The Problem of Hidden Variables in Quantum Mechanics, J. Math. Mech. **17**, 59 (1967).

[3] J. S. Bell, On the Einstein-Podolsky-Rosen Paradox, Physics **1**, 195 (1964). Reprinted in Ref. [22], chap. 2.

[4] J. S. Bell, On the problem of hidden variables in quantum mechanics, Rev. Mod. Phys. **38**, 447 (1966); Reprinted in Ref. [22], chap. 1.

[5] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. 23, 880 (1969).

[6] A. Einstein, B. Podolsky, and N. Rosen, Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?, Phys. Rev. 47, 777 (1935).

[7] Indeed, in Ref. [22] (p. 157), Bell writes

> My own first paper on [the subject of Bell's Theorem] ... starts with a summary of the EPR argument from locality to deterministic hidden variables. But the commentators have almost universally reported that it begins with deterministic hidden variables.

Although Wiseman has disputed Bell's account of the role of determinism in his first paper[19], see Norsen's response[20].

[8] R. W. Spekkens, Contextuality for preparations, transformations, and unsharp measurements, Phys. Rev. A **71**, 052108 (2005).

[9] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, An experimental test of noncontextuality without unwarranted idealizations, arXiv:1505.06244 (quant-ph) (2015).

[10] A. Cabello, J. Estebaranz, and G. Garcia-Alcaine, Bell-Kochen-Specker theorem: A proof with 18 vectors, Physics Letters A **212**, 183 (1996).

[11] D. A. Meyer, Finite Precision Measurement Nullifies the Kochen-Specker Theorem, Phys. Rev. Lett. 83, 3751 (1999); A. Kent, Noncontextual Hidden Variables and Physical Measurements, Phys. Rev. Lett. 83, 3755 (1999); R. Clifton and A. Kent, Simulating quantum mechanics by non-contextual hidden variables, Proc. R. Soc. Lond. A: 2000 456 2101-2114 (2000); J. Barrett and A. Kent, Non-contextuality, finite precision measurement and the Kochen-Specker theorem, Stud. Hist. Philos. Mod. Phys. 35, 151 (2004).

[12] Indeed, in Ref. [21], David Mermin is quoted as having said: "the whole notion of an experimental test of [B]KS misses the point", a view that was held by many researchers at the time.

[13] A. Cabello, Experimentally testable state-independent quantum contextuality, Phys. Rev. Lett. **101**, 210401 (2008).

[14] M. F. Pusey, The robust noncontextuality inequalities in the simplest scenario (2015).

[15] R. W. Spekkens, Negativity and Contextuality are Equivalent Notions of Nonclassicality, Phys. Rev. Lett. **101**, 020401 (2008).

[16] A. Cabello, S. Severini, and A. Winter, Graph-Theoretic Approach to Quantum Correlations, Phys. Rev. Lett.

112, 040401 (2014).
[17] A. Acin, T. Fritz, A. Leverrier, and A. B. Sainz, A Combinatorial Approach to Nonlocality and Contextuality, Comm. Math. Phys. 334(2), 533-628 (2015).
[18] R. W. Spekkens, The Status of Determinism in Proofs of the Impossibility of a Noncontextual Model of Quantum Theory, Found. Phys. **44**, 1125 (2014).
[19] H. M. Wiseman, The Two Bell's theorems of John Bell, J. Phys. A: Math. Theor. 47 424001 (2014).
[20] T. Norsen, Are there really two different Bell's theorems, arXiv:1503.05017 (quant-ph) (2015).
[21] A. Cabello and G. Garcia-Alcaine, Proposed Experimental Tests of the Bell-Kochen-Specker Theorem, Phys. Rev. Lett. 80, 1797 (1998).
[22] J. S. Bell, *Speakable and unspeakable in quantum mechanics* (Cambridge University Press, New York, 1987).

Appendix A: Proof of the inequality

We can summarize our main result—a derivation of a noncontextuality inequality from the proof of the Kochen-Specker theorem for the 18 ray uncolourable set of Fig. 1—by the following theorem:

Theorem. *Consider an operational theory* $(\mathcal{P}, \mathcal{M}, p)$. *Let* $\{M_i \in \mathcal{M} : i \in \{1, \ldots, 9\}\}$ *be nine four-outcome measurements. Let* $[k|M_i]$ *denote the kth outcome of the ith measurement, where* $k \in \{1, \ldots, 4\}$. *Let* $\{P_{i,k} \in \mathcal{P} : i \in \{1, \ldots, 9\},\ k \in \{1, 2, 3, 4\}\}$ *be thirty-six preparation procedures, organized into nine sets of four. Let* $P_i^{(\mathrm{ave})} \in \mathcal{P}$ *be the preparation procedure obtained by sampling* $k \in \{1, 2, 3, 4\}$ *uniformly at random and implementing* $P_{i,k}$.

Suppose that one has experimentally verified the operational preparation equivalences depicted in Fig. 3, namely,

$$P_1^{(\mathrm{ave})} \simeq P_2^{(\mathrm{ave})} \simeq \cdots \simeq P_9^{(\mathrm{ave})}, \tag{A1}$$

and the operational equivalences depicted in Fig. 2(a), namely,

$$[k|M_i] \simeq [k'|M_{i'}], \tag{A2}$$

for the eighteen pairs specifed therein.

If one assumes that the operational theory admits of a universally noncontextual ontological model, that is, one which is both measurement-noncontextual and preparation-noncontextual, then the following inequality on operational probabilities holds

$$A \equiv \frac{1}{36} \sum_{i=1}^{9} \sum_{k=1}^{4} p(k|M_i, P_{i,k}) \leq \frac{5}{6}. \tag{A3}$$

We now provide the proof. For clarity, we expand on some of the steps presented in the main article.

Using Eq. (1), the quantity A can be expressed in terms of the distributions and response functions of the ontological model as

$$A = \frac{1}{36} \sum_{i=1}^{9} \sum_{k=1}^{4} \sum_{\lambda} \xi(k|M_i, \lambda) \mu(\lambda|P_{i,k}). \tag{A4}$$

Using the definition of the max-probability $\zeta(M_i, \lambda)$, given in Eq. (12), we have

$$A \leq \frac{1}{9} \sum_{i=1}^{9} \sum_{\lambda} \zeta(M_i, \lambda) \left(\frac{1}{4} \sum_{k=1}^{4} \mu(\lambda|P_{i,k}) \right). \tag{A5}$$

Assuming that one experimentally verifies the operational preparation equivalences of Eq. (A1), the assumption of preparation noncontextuality implies that

$$\mu(\lambda|P_1^{(\mathrm{ave})}) = \mu(\lambda|P_2^{(\mathrm{ave})}) = \cdots = \mu(\lambda|P_9^{(\mathrm{ave})}). \tag{A6}$$

It follows that there exists a single distribution, which we denote $\nu(\lambda)$, such that

$$\mu(\lambda|P_i^{(\mathrm{ave})}) = \nu(\lambda) \ \text{ for all } i \in \{1, \ldots, 9\}. \tag{A7}$$

Recall that $P_i^{(\mathrm{ave})}$ is the preparation procedure that samples k uniformly from $\{1, 2, 3, 4\}$ and implements $P_{i,k}$. Given that the probability of the system being in a given ontic state λ given the preparation $P_{i,k}$ is $\mu(\lambda|P_{i,k})$, and given that the probability of $P_{i,k}$ being implemented is $\frac{1}{4}$ for each value of k, it follows that the probability of the system being in a given ontic state λ given the preparation $P_i^{(\mathrm{ave})}$ is $\mu(\lambda|P_i^{(\mathrm{ave})}) = \frac{1}{4} \sum_{\lambda} \mu(\lambda|P_{i,k})$. Combining this with Eq. (A7), we conclude that

$$\frac{1}{4} \sum_{\lambda} \mu(\lambda|P_{i,k}) = \nu(\lambda) \ \text{ for all } i \in \{1, \ldots, 9\}, \tag{A8}$$

and therefore that

$$A \leq \frac{1}{9} \sum_{\lambda} \sum_{i=1}^{9} \zeta(M_i, \lambda) \nu(\lambda). \tag{A9}$$

This in turn implies

$$A \leq \max_{\lambda} \frac{1}{9} \sum_{i=1}^{9} \zeta(M_i, \lambda). \tag{A10}$$

Assuming that one experimentally verifies the operational measurement equivalences of Eq. (A2), the assumption of measurement noncontextuality implies that

$$\xi(k|M_i, \lambda) = \xi(k'|M_{i'}, \lambda), \tag{A11}$$

for the eighteen pairs of operationally equivalent measurement events $([k|M_i], [k'|M_{i'}])$ specifed in Fig. 2(a).

It is useful to simplify the notation at this stage. We introduce the variable $\kappa \in \{1, \ldots, 18\}$ to range over the eighteen operational equivalence classes of measurement events. We introduce the shorthand notation

$$w_\kappa \equiv \xi(k|M_i, \lambda) = \xi(k'|M_{i'}, \lambda), \tag{A12}$$

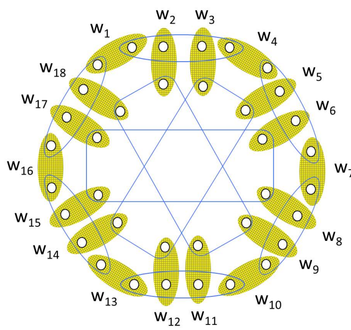

FIG. 5. A choice of labelling of the eighteen equivalence classes of measurement events. Here, w_κ denotes the probability assigned to the equivalence class labelled by κ in a noncontextual outcome-indeterministic ontological model.

for the probability assigned to the κth equivalence class, where the dependence on λ is left implicit. The variable κ enumerates the equivalence classes in Fig. 2(a) starting from $[1|M_1]$ and proceeding clockwise around the hypergraph, as depicted in Fig. 5.

In this notation, the constraint that each response function is probability-valued, $\xi(k|M_i.\lambda) \in [0,1]$, is simply

$$0 \leq w_\kappa \leq 1, \;\; \forall \kappa \in \{1,\ldots,18\}, \tag{A13}$$

while the constraint that the set of response functions for each measurement sum to 1, $\sum_{k=1}^{4} \xi(k|M_i,\lambda) = 1$, can be captured by the matrix equality

$$Z\vec{w} = \vec{u} \tag{A14}$$

where $\vec{w} \equiv (w_1,\ldots,w_{18})^T$, $\vec{u} \equiv (1,1,1,1,1,1,1,1,1)^T$, and

$$Z \equiv \begin{pmatrix} 1&1&1&1&0&0&0&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&1&1&1&1&0&0&0&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&1&1&1&1&0&0&0&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&1&1&1&1&0&0&0&0&0 \\ 0&0&0&0&0&0&0&0&0&0&0&0&1&1&1&1&0&0 \\ 1&0&0&0&0&0&0&0&0&0&0&0&0&0&0&1&1&1 \\ 0&1&0&0&0&0&0&0&1&0&1&0&0&0&0&0&0&1 \\ 0&0&1&0&1&0&0&0&0&0&0&1&0&1&0&0&0&0 \\ 0&0&0&0&0&1&0&1&0&0&0&0&0&0&1&0&1&0 \end{pmatrix}. \tag{A15}$$

Finally, we can express the quantity to be maximized as

$$\frac{1}{9}\sum_{i=1}^{9} \zeta(M_i,\lambda) = \frac{1}{9}\sum_{i=1}^{9} \max_{\kappa: Z_{i\kappa}=1} w_\kappa, \tag{A16}$$

or, more explicitly, as

$$\begin{aligned} &\frac{1}{9}\sum_{i=1}^{9} \zeta(M_i,\lambda) \\ &= \frac{1}{9}[\max\{w_1,w_2,w_3,w_4\} + \max\{w_4,w_5,w_6,w_7\} \\ &+ \max\{w_7,w_8,w_9,w_{10}\} + \max\{w_{10},w_{11},w_{12},w_{13}\} \\ &+ \max\{w_{13},w_{14},w_{15},w_{16}\} + \max\{w_{16},w_{17},w_{18},w_1\} \\ &+ \max\{w_{18},w_2,w_9,w_{11}\} + \max\{w_3,w_5,w_{12},w_{14}\} \\ &+ \max\{w_6,w_8,w_{15},w_{17}\}]. \end{aligned} \tag{A17}$$

The matrix equality of Eq. (A14) implies that there are only nine independent variables in the set $\{w_1,w_2,\ldots,w_{18}\}$ and that these satisfy linear inequalities. The space of possibilities for the vector $\vec{w}$ therefore forms a nine-dimensional polytope in the hypercube described by Eq. (A13).

The value of $\frac{1}{9}\sum_{i=1}^{9}\zeta(M_i,\lambda)$ on any of the interior points of this polytope will be an average of its values at the vertices because it is a convex function of $\vec{w}$. Therefore, to implement the maximization over λ, it suffices to maximize over the vertices of this polytope.

Following a brute-force enumeration of all the vertices of the polytope, the maximum possible value of $\frac{1}{9}\sum_{i=1}^{9}\zeta(M_i,\lambda)$ is found to be $\frac{5}{6}$. An example of a vertex achieving this value is $\vec{w} = (1,0,0,0,1,0,0,0,\frac{1}{2},\frac{1}{2},\frac{1}{2},0,0,0,1,0,0,0)^{\mathrm{T}}$, which is depicted in Fig. 4. This concludes the proof.

Our proof technique can be adapted to derive a similar noncontextuality inequality correponding to any proof of the KS theorem based on the uncolourability of a set of rays of Hilbert space. One begins by completing every set of orthogonal rays into a basis of the Hilbert space, and then forming the hypergraph depicting the orthogonality relations among these rays (the analogue of Fig. 1). One then forms the hypergraph decipting all of the measurements events, with one type of edge denoting which events correspond to the outcomes of a single measurement, and the other type of edge denoting when a set of measurement events are operationally equivalent (the analogue of Fig. 2(a)). One then associates a set of preparations with every measurement in the hypergraph, one preparation for every outcome. For each such set of preparations, we define the effective preparation that is the uniform mixture of the set's elements, and we presume that all of the effective preparations so defined are operationally equivalent (as is the case in quantum theory, where the effective preparation for every set corresponds to the completely mixed state). We consider the correlation between the measurement outcome and the choice of preparation in the set associated with that measurement, averaged over all measurements. This average correlation is the quantity A that appears on the left-hand side of the operational inequality.

The uncolourability of the hypergraph means that there are no noncontextual deterministic assignments to the measurement events, hence the polytope of proba-

bilistic assignments to the measurement events has no deterministic vertices either. Each vertex of this polytope, that is, each convexly-extremal probabilistic assignment, will necessarily yield an indeterministic assignment to some of the measurement events. Using the operational equivalences and the assumption of universal noncontextuality, one can infer from this that the average correlation A is always bounded away from 1. For any uncolourable hypergraph, a quantum realization would achieve the logical limit $A = 1$ by construction, so the noncontextuality inequality we derive is necessarily violated by quantum theory in each case.

One can understand this violation as being due to the fact that assignments of density operators that are independent of the preparation context can achieve higher predictability for the respective measurements than assignments of probability distributions over ontic states that are independent of the preparation context. This is the feature of quantum theory that allows it to maximally violate the noncontextual bound of $A \leq 5/6$.

Appendix B: Robustness of the noncontextuality inequality to noise

How much noise can one add to the measurements and preparations while still violating our noncontextuality inequality? We answer this question here assuming that the experimental operations are well-modelled by quantum theory. According to quantum theory,

$$p(k|M_i, P_{i,k}) = \mathrm{Tr}(E_{k|M_i}\rho_{i,k}), \tag{B1}$$

where $E_{k|M_i}$ denotes the positive operator representing the measurement event $[k|M_i]$ and $\rho_{i,k}$ denotes the density operator representing the preparation $P_{i,k}$. To be precise, for every i, the set $\{E_{k|M_i}\}_k$ is a positive operator valued measure, so that $0 \leq E_{k|M_i} \leq I$, and $\sum_k E_{k|M_i} = I$, and for every i and k, $\rho_{i,k}$ is positive, $\rho_{i,k} \geq 0$, and has unit trace, $\mathrm{Tr}\rho_{i,k} = 1$.

In quantum theory, a noiseless and maximally informative measurement is represented by a POVM whose elements are rank-1 projectors, that is,

$$E_{k|M_i} = \Pi_{i,k}, \tag{B2}$$

where for each k, $\Pi_{i,k}$ is a projector, hence idempotent, $\Pi_{i,k}^2 = \Pi_{i,k}$, and is rank 1, so that $\Pi_{i,k} = |\psi_{i,k}\rangle\langle\psi_{i,k}|$, where for each i, the set $\{|\psi_{i,k}\rangle\}_k$ is an orthonormal basis of the Hilbert space. If we furthermore set

$$\rho_{i,k} = \Pi_{i,k}, \tag{B3}$$

then we find $p(k|M_i, P_{i,k}) = \mathrm{Tr}(E_{k|M_i}\rho_{i,k}) = 1$ for each (i,k), and consequently $A = 1$. We see, therefore, that the maximum possible value of A is attained when measurements satisfy the noiseless ideal. We can now consider the consequence of adding noise.

We begin by considering a very simple noise model wherein the preparations and measurements both deviate from the noiseless ideal by the action of a depolarizing channel, that is, a channel of the form

$$\mathcal{D}_p(\cdot) = pI(\cdot)I + (1-p)\frac{1}{4}I\ \mathrm{Tr}(\cdot), \tag{B4}$$

which with probability p implements the identity channel and with probability $1-p$ generates the completely mixed state. If the quantum states are the image of the ideal states under a depolarizing channel with parameter p_1, and the POVM is obtained by acting the depolarizing channel with parameter p_2 followed by the ideal projector-valued measure (such that the POVM elements are the images of the projectors under the *adjoint* of the channel), then

$$\rho_{i,k} = \mathcal{D}_{p_1}(\Pi_{i,k}) = p_1\Pi_{i,k} + (1-p_1)\frac{1}{4}I, \tag{B5}$$

$$E_{k|M_i} = \mathcal{D}^\dagger_{p_2}(\Pi_{i,k}) = p_2\Pi_{i,k} + (1-p_2)\frac{1}{4}I, \tag{B6}$$

Here, the POVM $\{E_{k|M_i}\}_k$ is a mixture of $\{\Pi_{i,k}\}_k$ and a POVM $\{\frac{1}{4}I, \frac{1}{4}I, \frac{1}{4}I, \frac{1}{4}I\}$ which simply samples k uniformly at random regardless of the input state. It follows that for each (i,k), if we consider $p(k|M_i, P_{i,k}) = \mathrm{Tr}(E_{k|M_i}\rho_{i,k})$, we find perfect predictability for the term having weight p_1p_2 while for the three other terms, we have a uniformly random outcome, so that in all

$$p(k|M_i, P_{i,k}) = p_1p_2 + (1-p_1p_2)\frac{1}{4}. \tag{B7}$$

It follows that

$$A \equiv \frac{1}{36}\sum_{i=1}^{9}\sum_{k=1}^{4} p(k|M_i, P_{i,k}) = \frac{1}{4} + \frac{3}{4}p_1p_2, \tag{B8}$$

Thus a violation of the noncontextuality inequality, i.e. $A > \frac{5}{6}$, occurs if and only if

$$p_1p_2 > \frac{7}{9}. \tag{B9}$$

It turns out that one can derive similar bounds for more general noise models as well. Suppose that instead of a depolarizing channel, we have one of the form

$$\mathcal{N}_{p,\rho}(\cdot) = pI(\cdot)I + (1-p)\rho\ \mathrm{Tr}(\cdot). \tag{B10}$$

With probability p, this implements the identity channel and with probability $1-p$ it reprepares a state ρ that need not be the completely mixed state, but which is independent of the input to the channel. The analogous sort of noise acting on the measurement corresponds to acting on the POVM elements by the adjoint of this channel, that is,

$$\mathcal{N}^\dagger_{p,\rho}(\cdot) = pI(\cdot)I + (1-p)I\ \mathrm{Tr}(\rho\ \cdot). \tag{B11}$$

Therefore, if this sort of noise is applied to the ideal states and measurements, with the parameters in each

noise model allowed to depend on i, we obtain

$$\rho_{i,k} = \mathcal{N}_{p_1^{(i)},\rho_i}(\Pi_{i,k}) = p_1^{(i)}\Pi_{i,k} + (1-p_1^{(i)})\rho^{(i)}, \text{(B12)}$$

$$E_{k|M_i} = \mathcal{N}^{\dagger}_{p_2^{(i)},\sigma_i}(\Pi_{i,k}) = p_2^{(i)}\Pi_{i,k} + (1-p_2^{(i)})s(k|i)I, \tag{B13}$$

where $s(k|i) \equiv \mathrm{Tr}(\rho^{(i)}\Pi_{i,k})$ is a probability distribution over k for each value of i. Here, the POVM $\{E_{k|M_i}\}_k$ is a mixture of $\{\Pi_{i,k}\}_k$ and a POVM $\{s(k|i)I\}_k$ which simply samples k at random from the distribution $s(k|i)$, regardless of the quantum state. Compared to the simple model considered above, the innovation of this one is that for both preparations and measurements, the noise is allowed to be biased.

For the case of $p_1^{(i)} = 0$, which by Eq. (B12) implies that $\rho_{i,k} = \rho^{(i)}$, we find that, regardless of the measurement, $p(k|M_i, P_{i,k})$ is just a normalized probability distribution over k (because there is no k dependence in the state). Hence, in this case, $\frac{1}{4}\sum_{k=1}^{4} p(k|M_i, P_{i,k}) = \frac{1}{4}$.

Similarly, for the case of $p_2^{(i)} = 0$, that is, when the POVM corresponds to a random number generator $E_{k|M_i} = s(k|i)I$, we find that, regardless of the preparation, $p(k|M_i, P_{i,k})$ is again just a normalized probability distribution over k. Hence, in this case again, $\frac{1}{4}\sum_{k=1}^{4} p(k|M_i, P_{i,k}) = \frac{1}{4}$.

It follows that for generic values of $p_1^{(i)}$ and $p_2^{(i)}$, we have $\frac{1}{4}\sum_{k=1}^{4} p(k|M_i, P_{i,k}) = p_1^{(i)}p_2^{(i)} + (1 - p_1^{(i)}p_2^{(i)})\frac{1}{4}$. In all then, we have

$$A \equiv \frac{1}{36}\sum_{i=1}^{9}\sum_{k=1}^{4} p(k|M_i, P_{i,k}) = \frac{1}{4} + \frac{3}{4}\left(\frac{1}{9}\sum_{i=1}^{9} p_1^{(i)}p_2^{(i)}\right). \tag{B14}$$

Consequently, a violation of the noncontextuality inequality, i.e., $A > \frac{5}{6}$, occurs if and only if the noise parameters satisfy

$$\frac{1}{9}\sum_{i=1}^{9} p_1^{(i)}p_2^{(i)} > \frac{7}{9}. \tag{B15}$$

Because the parameters $p_1^{(i)}$ and $p_2^{(i)}$ decrease as one increases the amount of noise, this inequality specifies an upper bound on the amount of noise that can be tolerated if one seeks to violate the noncontextuality inequality.

This analysis highlights how the approach to deriving noncontextuality inequalities described in this article has no trouble accommodating noisy POVMs. This contrasts with previous proposals for experimental tests based on the traditional notion of noncontextuality, which can only be applied to projective measurements. This is one way to see how previous proposals are not applicable to realistic experiments, where every measurement has some noise and consequently is necessarily *not* represented projectively.

Appendix C: Comparison to other noncontextuality inequalities

We have proposed a technique for deriving noncontextuality inequalities from proofs of the Kochen-Specker theorem. It is useful to compare our approach with one that has previously been proposed by Cabello [13]. We do so by explicitly comparing the two proposals in the case of the 18 ray construction of Ref. [10]. Indeed, the fact that Ref. [13] proposes an inequality for this construction is part of our motivation for choosing it as our illustrative example.

For each of the eighteen operational equivalence classes of measurement events, labelled by $\kappa \in \{1, \ldots, 18\}$ as depicted in Fig. 5, we associate a $\{-1, +1\}$-valued variable, denoted $S_\kappa \in \{-1, +1\}$. A given ontic state λ is assumed to assign a value to each S_κ. The fact that there is only a *single* variable associated to each equivalence class implies that any assignment of such values is necessarily noncontextual.

Ref. [13] considers a particular linear combination of expectation values of products of these variables:

$$\begin{aligned}\alpha \equiv & -\langle S_1S_2S_3S_4\rangle - \langle S_4S_5S_6S_7\rangle - \langle S_7S_8S_9S_{10}\rangle \\ & - \langle S_{10}S_{11}S_{12}S_{13}\rangle - \langle S_{13}S_{14}S_{15}S_{16}\rangle - \langle S_{16}S_{17}S_{18}S_1\rangle \\ & - \langle S_{18}S_2S_9S_{11}\rangle - \langle S_3S_5S_{12}S_{14}\rangle \\ & - \langle S_6S_8S_{15}S_{17}\rangle, \end{aligned} \tag{C1}$$

and derives the following inequality for it:

$$\alpha \le 7 \tag{C2}$$

(Note that Ref. [13] used a labelling convention for the eighteen measurement events that is different from the one we use here; to translate between the two conventions, it suffices to compare Fig. 1 in that article with Fig. 5 in ours.) Each term in α refers to a quadruple of variables that can be measured together, that is, which can be computed from the outcome of a single measurement. Different terms correspond to measurements that are incompatible.

In Ref. [13], the following justification is given for the inequality (C2). We are asked to consider the 2^{18} possible assignments to $(S_1, \ldots, S_{18})$ that result from the two possible assignments to S_κ, namely -1 or $+1$, for each $\kappa \in \{1, \ldots, 18\}$. It is then noted that among all such possibilities, the maximum value of α that can be achieved is 7.

Ref. [13] states that a violation of this inequality should be considered evidence of a failure of noncontextuality. We disagree with this conclusion, and the rest of this section seeks to explain why.

1. The most natural interpretation

It is useful to recast the inequality of Eq. (C2) in terms of variables v_κ with values in $\{0, 1\}$ rather than $\{-1, +1\}$.

Specifically, we take

$$v_\kappa \equiv \frac{S_\kappa + 1}{2}. \tag{C3}$$

Under this translation, products of the S_κ correspond to sums (modulo 2) of the v_κ. For instance, an equation such as $S_{\kappa_1}S_{\kappa_2} = -1$ corresponds to the equation $v_{\kappa_1} \oplus v_{\kappa_2} = 1$, where $\oplus$ denotes sum modulo 2, while $S_{\kappa_1}S_{\kappa_2} = +1$ corresponds to $v_{\kappa_1} \oplus v_{\kappa_2} = 0$, so that $v_{\kappa_1} \oplus v_{\kappa_2} = \frac{-S_{\kappa_1}S_{\kappa_2}+1}{2}$. In particular, we also have

$$v_{\kappa_1} \oplus v_{\kappa_2} \oplus v_{\kappa_3} \oplus v_{\kappa_4} = \frac{-S_{\kappa_1}S_{\kappa_2}S_{\kappa_3}S_{\kappa_4} + 1}{2} \tag{C4}$$

or equivalently,

$$-S_{\kappa_1}S_{\kappa_2}S_{\kappa_3}S_{\kappa_4} = 2(v_{\kappa_1} \oplus v_{\kappa_2} \oplus v_{\kappa_3} \oplus v_{\kappa_4}) - 1, \tag{C5}$$

We can therefore consider a quantity α', defined as

$$\begin{aligned}\alpha' \equiv &\langle v_1 \oplus v_2 \oplus v_3 \oplus v_4\rangle + \langle v_4 \oplus v_5 \oplus v_6 \oplus v_7\rangle \\ &+ \langle v_7 \oplus v_8 \oplus v_9 \oplus v_{10}\rangle + \langle v_{10} \oplus v_{11} \oplus v_{12} \oplus v_{13}\rangle \\ &+ \langle v_{13} \oplus v_{14} \oplus v_{15} \oplus v_{16}\rangle + \langle v_{16} \oplus v_{17} \oplus v_{18} \oplus v_1\rangle \\ &+ \langle v_{18} \oplus v_2 \oplus v_9 \oplus v_{11}\rangle + \langle v_3 \oplus v_5 \oplus v_{12} \oplus v_{14}\rangle \\ &+ \langle v_6 \oplus v_8 \oplus v_{15} \oplus v_{17}\rangle,\end{aligned} \tag{C6}$$

so that $\alpha = 2\alpha' - 9$, and we can re-express inequality (C2) as

$$\alpha' \leq 8. \tag{C7}$$

Of course, rather than using Eq. (C5) to translate (C2) from $\{-1,+1\}$-valued variables into $\{0,1\}$-valued variables, one can also just derive the inequality (C7) directly: among the 2^{18} possible assignments of values in $\{0,1\}$ to each of the v_κ, the maximum value of α' is 8. Two examples of such assignments are provided in Fig. 6.

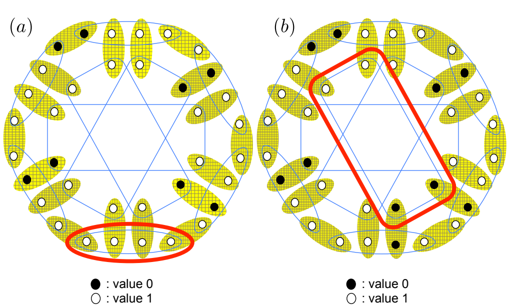

FIG. 6. Examples of noncontextual assignments of $\{0,1\}$-values to the measurement events in Fig. 2(a) where it is not required that every measurement has precisely one outcome that is assigned value 1 and three outcomes that are assigned the value 0. Example (a) depicts an assignment wherein there is a measurement all of whose outcomes receive probability 0. Example (b) depicts one wherein there is a measurement two of whose outcomes recieve probability 1.

It is useful to use a notation that specfies whether a given expectation value of some variable X is relative to a preparation procedure P, in which case it is denoted $\langle X\rangle_P$, or relative to an ontic state λ, in which case it is denoted $\langle X\rangle_\lambda$. We denote by $\alpha'(P)$ the quantity defined in (C6) if the expectation values contained therein are relative to preparation P, and we denote by $\alpha'(\lambda)$ the case where the expectation values are relative to ontic state λ. Under the assumption of an ontological model, each expectation value relative to a preparation P can be expressed as a function of the expectation value relative to an ontic state λ, via

$$\langle X\rangle_P = \sum_\lambda \langle X\rangle_\lambda \, \mu(\lambda|P), \tag{C8}$$

where $\mu(\lambda|P)$ is the distribution over ontic states associated with preparation P. We can infer from Eq. (C8) that

$$\alpha'(P) = \sum_\lambda \alpha'(\lambda)\mu(\lambda|P). \tag{C9}$$

With these notational conventions, we can summarize the argument of Ref. [13] as follows. In any noncontextual ontological model, every ontic state λ satisfies

$$\alpha'(\lambda) \leq 8. \tag{C10}$$

But this in turn implies, through Eq. (C9), that for all preparations P,

$$\alpha'(P) \leq 8, \tag{C11}$$

which is an inequality constraining operational quantities.

We are now in a position to describe the problem with the inequality (C11), or equivalently inequality (C2), and thus with the claim of Ref. [13]. First, we highlight the physical interpretation of the variables v_κ. If v_κ is assigned value 1 by the ontic state λ, then this means that if the system is in the ontic state λ, and a measurement that includes κ as an outcome is implemented on it, then the outcome κ is certain to occur, while if v_κ is assigned value 0 by λ, then the outcome κ is certain *not* to occur. But each of the 2^{18} different assignments to $(v_1,\ldots,v_{18})$ is such that for at least one measurement either: *none* of the outcomes occur, as in the example of Fig. 6(a), or *more than one* outcome occurs, as in the example of Fig. 6(b). (This is precisely what is implied by the fact that the 18 measurement events are *uncolourable*, as explained in the main text.) Such assignments involve a *logical contradiction* given that the four outcomes of each measurement are mutually excusive and jointly exhaustive possibilities.

It follows that the sort of model that a violation of inequality (C11) rules out can already be ruled out *by logic alone*; no experiment is required. To put it another way, discovering that quantum theory and nature violate inequality (C11) only allows one to conclude that neither quantum theory nor nature involve a logical contradiction, which one presumably already knew prior to noting the violation.

We have argued in the main text that the notion of KS-noncontextuality, insofar as it assumes outcome-determinism, is not suitable for devising experimentally robust inequalities given that every real measurement involves some noise. The problem with inequality (C11) can also be traced back to the use of the assumption of KS-noncontextuality. Suppose we ask the following question: given the existence of nine four-outcome measurements satisfying the operational equivalences of Fig. 2(a), how are the operational probabilities that are assigned to these measurement events constrained if we presume that KS-noncontextual assignments underlie the operational statistics? On the face of it, the question seems well-posed. On further reflection, however, one sees that it is not. There are simply *no* KS-noncontextual assignments to these measurement events, so it is simply impossible to imagine that such assignments could underlie the operational statistics. There is nothing to be tested experimentally, as the hypothesis under consideration is seen to be false as a matter of logic.

Here is another way to see that the inequality (C11) does not provide a test of noncontextuality. Consider the expectation value $\langle v_{\kappa_1} \oplus v_{\kappa_2} \oplus v_{\kappa_3} \oplus v_{\kappa_4} \rangle_P$ for a preparation P, where κ_1, κ_2, κ_3 and κ_4 correspond to the four outcomes of some measurement. Regardless of which of the four outcomes of the measurement occurs in a given run where preparation P is implemented—i.e. regardless of whether $(v_{\kappa_1}, v_{\kappa_2}, v_{\kappa_3}, v_{\kappa_4})$ comes out as (1,0,0,0) or (0,1,0,0) or (0,0,1,0) or (0,0,0,1) in that run—the variable $v_{\kappa_1} \oplus v_{\kappa_2} \oplus v_{\kappa_3} \oplus v_{\kappa_4}$ has the value 1. We can think of it this way: the variable $v_{\kappa_1} \oplus v_{\kappa_2} \oplus v_{\kappa_3} \oplus v_{\kappa_4}$ is a trivial variable because it is a constant function of the measurement outcome. (This is analogous to how, in quantum theory, for a four-outcome measurement associated with four projectors, although each projector is a nontrivial observable, their sum is the identity operator, which has expectation value 1 for all quantum states, and therefore corresponds to a trivial observable.) It follows that regardless of what distribution over the four outcomes is assigned by P, the expectation value $\langle v_{\kappa_1} \oplus v_{\kappa_2} \oplus v_{\kappa_3} \oplus v_{\kappa_4} \rangle_P$ will be 1. Given that each of the nine terms in $\alpha'(P)$ is of this form, it follows that $\alpha'(P) = 9$.

So, for *any* operational theory that admits of nine four-outcome measurements with the operational equivalence relations depicted in Fig. 2(a), we will find that $\alpha'(P) = 9$ for all P. Therefore, we can conclude that the inequality $\alpha'(P) \leq 8$ is violated for all P. One can reach this conclusion without ever considering the question of whether the operational predictions can be explained by some underlying noncontextual model.

Another consequence of the triviality of the variables of the form $v_{\kappa_1} \oplus v_{\kappa_2} \oplus v_{\kappa_3} \oplus v_{\kappa_4}$ is that the inequality (C11) can be violated regardless of how noisy the measurements are. Suppose, for instance, that quantum theory describes our experiment, but that the nine four-outcome measurements are not the projective measurements described in Fig. (1), but rather noisy versions thereof. For instance, one can imagine that each measurement is associated with a positive operator-valued measure that is the image under a depolarizing map of the projector valued measure associated with the ideal measurement. The amount of depolarization can be taken arbitrarily large and, as long as it is the *same* amount of depolarization for each of the measurements, the nine noisy measurements that result will still satisfy precisely the same operational equivalences as the original nine, namely, those depicted in Fig. 2(a). For such noisy measurements, we can still identify variables v_κ associated to the eighteen equivalence classes of measurement events, and we still find that regardless of which of the four outcomes of the measurement occurs, the variable $v_{\kappa_1} \oplus v_{\kappa_2} \oplus v_{\kappa_3} \oplus v_{\kappa_4}$ has the value 1, so that regardless of what distribution over the four outcomes is assigned by P, the expectation value $\langle v_{\kappa_1} \oplus v_{\kappa_2} \oplus v_{\kappa_3} \oplus v_{\kappa_4} \rangle_P$ will be 1 and therefore $\alpha'(P) = 9$, which is a violation of the inequality (C11).

According to the generalized notion of noncontextuality proposed in Ref. [8], if one adds enough noise to the preparations and measurements in an experiment, it always becomes possible to represent the experimental statistics by a noncontextual model. One way to prove this is to note that: (i) if all of the preparations and the measurements in an experiment admit of positive Wigner representations, then, as demonstrated in Ref. [15], the Wigner representation defines a noncontextual model, and (ii) if one adds enough noise to the preparations and measurements, it is possible to ensure that they admit of positive Wigner representations.

This analysis of the effect of noise accords with intuition: noncontextuality is meant to represent a notion of classicality, so that a failure of noncontextuality is only expected to occur in a quantum experiment if one's experimental operations have a high degree of coherence. It follows that there should always exist a threshold of noise above which an experiment cannot be used to demonstrate the failure of noncontextuality. One can turn this observation into a minimal criterion that should be satisfied by any noncontextuality inequality: there should exist a threshold of experimental noise above which a noncontextuality inequality cannot be violated.

As we have just noted, the inequality proposed in Ref. [13] fails this minimal criterion. By contrast, the noncontextuality inequality proposed in this article identifies such a threshold for the 18 ray construction: the noise must be kept low enough that the average of the measurement predictabilities is above 5/6.

2. Alternative interpretation

The inequality proposed in Ref. [13] can be given a different interpretation to the one provided in the previous subsection. This interpretation is more charitable in some ways, but it still does not vindicate the proposed inequality as delimiting the boundary of noncontextual models.

The idea is to imagine that for each of the nine mea-

surements, there are in fact *five* rather than four outcomes that are mutually exclusive and jointly exhaustive. Thus, in this interpretation, it is assumed that the hypergraph describing compatibility relations and operational equivalences is *not* the one of Fig. 2(a), but rather a modification wherein there are nine additional nodes—one additional node appended to each of the nine measurements—as depicted in Fig. 7(a).

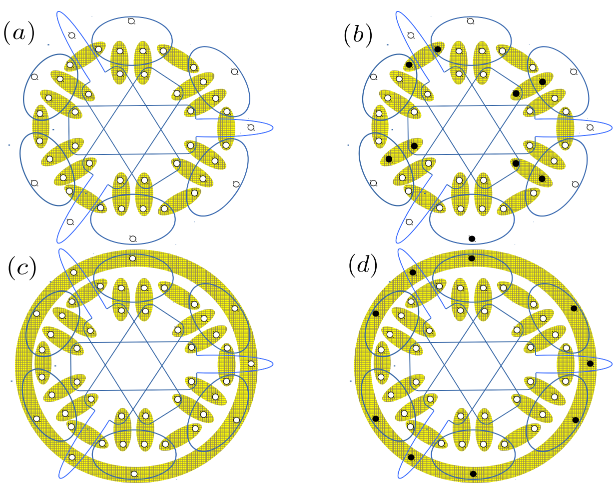

FIG. 7. (a) The hypergraph wherein each measurement is assigned an additional fifth outcome. (b) A normalized noncontextual deterministic assignment to the hypergraph of (a) that recovers the subnormalized noncontextual deterministic assignment of Fig. 6(a) on the appropriate subgraph; (c) The hypergraph wherein the fifth outcomes are all operationally equivalent; (d) the unique normalized noncontextual and deterministic assignment to the hypergraph of (c).

If $\{\kappa_1, \kappa_2, \kappa_3, \kappa_4\}$ are the original four outcomes of a given measurement, then the variable $v_{\kappa_1} \oplus v_{\kappa_2} \oplus v_{\kappa_3} \oplus v_{\kappa_4}$ is no longer a constant function of the measurement outcome because its value varies depending on whether or not the fifth outcome occurs. If κ_5 denotes the fifth outcome of the measurement, then the trivial variable is $v_{\kappa_1} \oplus v_{\kappa_2} \oplus v_{\kappa_3} \oplus v_{\kappa_4} \oplus v_{\kappa_5}$, taking the value 1 regardless of the outcome.

In this case, the assignments of the type depicted in Fig. 6(a)—the noncontextual deterministic assignments that are *subnormalized*—can be embedded into noncontextual deterministic *normalized* assignments on the larger hypergraph, as depicted in Fig. 7(b). (The possibility of such an embedding for the subnormalized noncontextual deterministic assignments considered in Cabello, Severini and Winter [16] was noted in Acin, Fritz, Leverrier, Sainz [17].)

Of course, such a move does not provide any way of understanding the deterministic noncontextual assignments of the type depicted in Fig. 6(b), because the latter violate normalization by having the probabilities of the different outcomes of the measurement summing to greater than 1—they are *supernormalized*.

So, while the supernormalized noncontextual deterministic assignments can be ruled out by logic alone, the subnormalized noncontextual deterministic assignments may be entertained without logical inconsistency if they are considered as reductions to a subgraph of a normalized noncontextual deterministic assignment on a larger hypergraph.

Because the justification given in Ref. [13] for the inequality derived there asks one to consider *all* of the noncontextual deterministic assignments, including the supernormalized ones, the interpretation of this inequality as a constraint on subnormalized assignments is in tension with the manner in which the inequality is justified. This interpretation is a better fit with Cabello's later work, such as Ref. [16], wherein the restriction to subnormalized assignments is explicit. In any case, if the inequality holds for *all* noncontextual deterministic assignments, regardless of normalization, then it holds for the special case of the subnormalized assignments, so the inequality can still be derived within this interpretation.

The problem with this interpretation becomes manifest when we require that the original hypergraph of Fig. 2(a)—and thus the corresponding subgraph of Fig. 7(a) from which it is derived in this interpretation—is realized in terms of Hilbert-space bases in the manner depicted in Fig. 1(a).

We consider two possible ways of fulfiling this requirement, and explain why it is not possible to vindicate the inequality of Eq. (C7) in either case.

In one approach, we imagine that the quantum system is in fact described by a 5-dimensional Hilbert space. In this case, rank-1 projective measurements have five outcomes and are therefore described within the hypergraph representation by an edge with five nodes, just as we have for the measurements in Fig. 7(a). Now consider an association of Hilbert space rays with the nodes of this hypergraph such that one recovers the association of rays to nodes described by Fig. 1 on the subgraph of Fig. 7(a) that corresponds to the original hypergraph of Fig. 2(a). This is possible if, for every measurement, the fifth outcome is associated with a ray that is orthogonal to the 4d subspace in which all of the other rays live. But then, under a tomographically complete set of preparations of the 5d Hilbert space, one finds that the fifth outcomes are all operationally equivalent, so that the appropriate hypergraph is not that of Fig. 7(a) but rather the one depicted in Fig. 7(c).

Now, consider *this* hypergraph. It only admits of a single normalized noncontextual deterministic assignment, the one that assigns 0s to every outcome in the original set and 1 to all of the fifth outcomes, as depicted in Fig. 7(d). Therefore, if one were to experimentally verify the applicability of the hypergraph of Fig. 7(c), by verifying the operational equivalences depicted therein, then any KS-noncontextual model consistent with this hypergraph would not only satisfy the inequality $\alpha'(\lambda) \leq 8$ (Eq. (C10)), it would predict that *all* of the measurement events appearing in the inequality receive value 0, so that the inequality could be strengthened to the

equality $\alpha'(\lambda) = 0$, which in turn would imply, through Eq. (C9), that for all preparation procedures P, the operational inequality $\alpha'(P) \leq 8$ could be strengthened to the operational equality

$$\alpha'(P) = 0. \tag{C12}$$

But this is trivial to violate experimentally: simply find a preparation that does not always yield the fifth outcome for every measurement.

We take the triviality of this constraint to speak against the idea that it captures the assumption of noncontextuality. Therefore, the conclusion to draw from this discussion is *not* that one should replace the inequality $\alpha'(P) \leq 8$ with $\alpha'(P) = 0$. Rather, as we've argued at length in the main text, because the KS-noncontextual models make the unjustified assumption of outcome-determinism, the notion of noncontextuality should not be formalized as KS-noncontextuality, but rather as measurement and preparation noncontextuality.

We now turn to the second approach. Here, one sticks to the notion that the quantum system being probed is 4-dimensional and instead one suggests that each of the nine measurements is nonprojective, that is, each is represented by a positive operator valued measure rather than a projector valued measure. In this way, one can ensure that the measurements indeed have five outcomes. One might even think of the fifth outcome as representing a 'no detection' event (the idea of justifying subnormalized assignments by imagining an additional 'no detection' outcome has also been discussed in Ref. [17]).

To see that there is something fishy about this approach, it suffices to note that if it were correct, then it would have the bizarre consequence that in the case where the measurements achieve the ideal of projectiveness, satisfaction of the inequality $\alpha'(P) \leq 8$ is ruled out by logic alone, whereas if the measurements depart from this ideal, however little, suddenly the inequality specifies whether or not the experiment can be modelled noncontextually.

In any case, the real problem with this approach is easily identified. For a nonprojective measurement, one is assigning probabilities to *effects* (positive operators less than identity) rather than projectors. In this case, one must allow noncontextual assignments to be *probabilistic*. This has been proven elsewhere [18] and we will not repeat the arguments here. Such probabilistic noncontextual assignments are not restricted to be in the convex hull of the deterministic noncontextual assignments, and therefore can be more general than mixtures of the latter. Because the derivation of the inequality $\alpha'(P) \leq 8$ made crucial use of the assumption that the preparation P was a mixture of deterministic noncontextual assignments, the fact that the assumption of determinism is unwarranted implies that one can no longer derive the inequality as a constraint on noncontextual models.